

AMENDMENTS TO THE CLAIMS

This listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

1. (Currently Amended) A computer implemented system that facilitates web-crawling, comprising at least a processor, one or more memories with following components stored thereon:
 - a managing component that performs a predictive analysis to predict when a web page will change, ~~in order to~~ and determines ~~if~~, when, and how to perform web-crawling;
 - a server computer component that implements a web-crawling component that crawls subsets of web pages as a function of the predictive analysis, ~~to discover and update~~ discovers and updates the pages in a catalogue of possible search results; and
 - a decision-theoretic component that determines an appropriate time to crawl the at least one web page and makes predictions regarding changes in at least one web page based at least in part on:
 - a probability that a particular outcome will occur, Pr; and
 - a utility factor associated with each outcome, Utility(O);
 - an action, a, selected from a set of possible actions, A, to be performed on the at least one web page, which maximizes the value of:

$$\sum_{o \in O} \Pr(o | a) \times \text{Utility}(o)$$

where o is an outcome selected from a set of possible outcomes, O, ~~so as to~~ wherein the outcome o maximizes ~~maximize~~ the efficiency of the web-crawling component in discovering and updating changed web pages.

- 2.-4. (Cancelled)

5. (Original) The system of claim 1, the predictive analysis is based at least in part on the utility of the at least one web page.

6. (Original) The system of claim 1, the predictive analysis is based at least in part on historical data related to the at least one web page.

7. (Original) The system of claim 1, the predictive analysis is based at least in part on content contained in the at least one web page.

8. (Previously Presented) The system of claim 1, further comprising a bundling component that rearranges crawled web pages into new subsets according to the utility of the web pages.

9. (Original) The system of claim 1, the web-crawling component comprises a Round Robin crawling component that sequentially crawls web pages in a subset and ensures that every web page will be crawled within a crawling period, and a Greedy crawling component that non-sequentially crawls pages according to a score associated with each page.

10.-37. (Cancelled)

38. (Currently Amended) A computer readable medium that has computer executable instructions stored thereon to:

predict when a web page will change in order to determine if, when, and how to perform web-crawling;

crawl subsets of web pages based on the predicting when a web page will change, to catalogue possible web page search results; and

determine an appropriate time to crawl the web page and make predictions regarding changes in at least one web page based at least in part on:

a probability that a particular outcome will occur, \Pr ; and

a utility factor associated with each outcome, $Utility(O)$;

an action, a , selected from a set of possible actions, A , to be performed on the at least one web page, which maximizes the value of:

$$\sum_{o \in O} \Pr(o | a) \times Utility(o)$$

where o is an outcome selected from a set of possible outcomes, O , so as to ~~wherein~~ wherein

the outcome o maximizes ~~maximize~~ the efficiency of crawling in discovering and updating changed web pages.

39.-41. (Cancelled)

42. (Previously Presented) The computer readable medium of claim 38, the instructions that predict are based at least in part on the utility of the at least one web page.

43. (Previously Presented) The computer readable medium of claim 38, the instructions that predict are based at least in part on historical data related to the at least one web page.

44. (Previously Presented) The computer readable medium of claim 38, the instructions that predict are based at least in part on content contained in the at least one web page.

45. (Previously Presented) The computer readable medium of claim 38, further comprising instructions that rearrange the crawled web pages into new subsets according to the utility of the web pages.

46. (Previously Presented) The computer readable medium of claim 38, the instructions that crawl comprise instructions that sequentially crawl web pages in a subset within a crawling period, and non-sequentially crawl pages according to a score associated with each page.

47.-48. (Cancelled)

49. (New) A computer readable medium having stored thereon components that facilitate web-crawling, the components comprising:

a managing component that performs a predictive analysis to predict when a web page will change, and determines when, and how to perform web-crawling;

a server computer component that implements a web-crawling component that crawls subsets of web pages as a function of the predictive analysis, discovers and updates the pages in a catalogue of possible search results; and

a decision-theoretic component that determines an appropriate time to crawl the at least one web page and makes predictions regarding changes in at least one web page based at least in part on:

a probability that a particular outcome will occur, Pr ; and

a utility factor associated with each outcome, $Utility(O)$;

an action, a , selected from a set of possible actions, A , to be performed on the at least one web page, which maximizes the value of:

$$\sum_{o \in O} Pr(o | a) \times Utility(o)$$

where o is an outcome selected from a set of possible outcomes, O , wherein the outcome o maximizes the efficiency of the web-crawling component in discovering and updating changed web pages.

50. (New) The medium of claim 49, the predictive analysis is based at least in part on the utility of the at least one web page.

51. (New) The medium of claim 49, the predictive analysis is based at least in part on historical data related to the at least one web page.

52. (New) The medium of claim 49, the predictive analysis is based at least in part on content contained in the at least one web page.

53. (New) The medium of claim 49, further comprising a bundling component that rearranges crawled web pages into new subsets according to the utility of the web pages.

54. (New) The medium of claim 49, the web-crawling component comprises a Round Robin crawling component that sequentially crawls web pages in a subset and ensures that every web page will be crawled within a crawling period, and a Greedy crawling component that non-sequentially crawls pages according to a score associated with each page.